Numerical Weather Prediction

An Assessment of Aircraft Icing Forecasts from the Met Office UK Mesoscale and Global Numerical Weather Prediction Models

Forecasting Research Technical Report No. 391

ian Macadam

email: nwp_publications@metoffice.com

©Crown Copyright

An Assessment of Aircraft Icing Forecasts from the Met Office UK Mesoscale and Global Numerical Weather Prediction Models

I. Macadam
Met Office

July 2002

Abstract

A set of experimental +6 to +18 hour aircraft icing forecasts from the Met Office UK mesoscale and global numerical weather prediction models are assessed. As well as assessing forecasts of areas of nonzero supercooled liquid water content (SLWC) from both models, equivalent forecasts of icing based on two types of icing intensity index from a new icing algorithm used by the Swedish Armed Forces Headquarters (SAFHQ) are also assessed. Hit rate is effectively assessed in terms of the horizontal and vertical separations between the locations of a set of 28 icing encounters reported by RAF pilots over the UK and the locations of the nearest forecast areas of potential icing. False alarm rate is effectively assessed in terms of the sizes of the areas for which potential icing conditions are forecast. The analysis suggests that the mesoscale model forecasts have a higher hit rate than the global model forecasts and that above 5000 ft they also have a false alarm rate approximately equal to or less than that of the global model forecasts. In term of hit rate, the nonzero SLWC forecasts from both models appear to perform better than the equivalent forecasts involving the SAFHQ icing intensity indices. However, it is possible that in terms of false alarm rate, the nonzero SLWC forecasts perform worse than the SAFHQ icing intensity index forecasts. Recommendations are made for a more thorough further study, which will be required before much more can be deduced about the relative performance of the different types of aircraft icing forecast discussed here.

1. Introduction

This is an initial assessment of +6 to +18 hour (hr) aircraft icing forecasts from the Met Office UK mesoscale and global numerical weather prediction (NWP) models [1], hereafter referred to as the mesoscale model and the global model respectively. Forecasts of between +0 and +6 hr were not investigated because it was thought unlikely that such short range forecasts would be used in practice.

The study described by this document assumes that areas of the atmosphere containing supercooled liquid water correspond to areas where aircraft might encounter icing. As well as assessing forecasts of areas of nonzero supercooled liquid water content (SLWC) from both models, corresponding forecasts of icing based on two types of icing intensity index from a new icing algorithm used by the Swedish Armed Forces Headquarters (SAFHQ) [2] are also assessed.

Section 2 of this report is a brief overview of the versions of the mesoscale and global models used in the study. Section 3 is an introduction to the new SAFHQ aircraft icing algorithm. Section 4 describes the basic method used to assess the forecasts from the models. Some general observations about the forecasts are made in Section 5. Sections 6 and 7 each discuss the performance of the forecasts with respect to one of the two main criteria important in the assessment of any set of aircraft icing forecasts. These criteria are:

(i) Hit Rate
This is a measure of how often the forecasts forecast icing when and where icing is actually reported. Obviously, for a set of forecasts to be useful it must have a high hit rate.

(ii) False Alarm Rate
Overprediction of icing is a common failing of icing forecasts that significantly reduces their value. Clearly, a set of forecasts that achieves a 100% hit rate by forecasting icing conditions everywhere at all times is worthless. False alarm rate is a measure of how often icing is forecast when and where observations indicate that there is actually no risk of icing. For a set of forecasts to be useful it must have a low false alarm rate.

In Section 8 conclusions are drawn in terms of these criteria and finally in Section 9 recommendations for further work in this area, based on the limitations of this study, are made.

2. The Models

The mesoscale and global models are not completely separate models. They are, in fact, two different configurations of the Met Office's Unified Model (UM). At the time of writing, May 2002, version 5.2 of the UM is about to replace version 4.5 as the source of all of the Met Office's operational mesoscale and global NWP forecasts. The mesoscale and global models used

in this assessment are configurations of different versions of the UM. The global model is a global NWP configuration of version 5.2, which, among other improvements, includes a new cloud scheme not included in version 4.5 of the global NWP configuration. The mesoscale model used in the study is a mesoscale NWP configuration of version 4.5, which also includes the new cloud scheme, but it is forced with boundary conditions from the global NWP configuration of version 4.5, which does not include the new cloud scheme. The observations on which this work relies were taken during 2000 and 2001, a period for which initial conditions for the global model consistent with version 5.2 were not readily available. Therefore, for the assessment of the global model the icing forecasts to be compared with the observations had to be obtained from forecast runs of version 5.2 of the model initialised with initial conditions derived from version 4.5 runs. As a consequence of this, it is possible that the performance of the global model forecasts presented in this study is not as good as the performance of forecasts from the Met Office operational global NWP model to be expected after version 5.2 of the UM becomes the standard operational version. It was not deemed necessary to obtain icing forecasts from version 5.2 of the mesoscale model because version 4.5 of this model already includes the new cloud scheme.

The mesoscale model used in this study has a domain that covers the British Isles and surrounding waters and extends into continental Europe as far as southwestern Sweden and the extreme northwest of Italy. The horizontal coordinate system is a rotated pole latitude-longitude grid with adjacent gridpoints approximately 6 nautical miles (nm) apart over the UK. There are 38 levels in the vertical, which follow constant values of a function of pressure, and the quantities used in this assessment are outputted on these levels. The mesoscale model is run operationally every 6 hr from initial conditions valid at 00:00, 06:00, 12:00 and 18:00.

The global model used in this study has a latitude-longitude grid horizontal coordinate system with adjacent gridpoints approximately 32 nm apart over the UK. Like the mesoscale model, the global model has 38 levels in the vertical but, unlike the mesoscale model, these follow constant values of a function of height. However, the model outputs the quantities used in this assessment on surfaces of constant potential temperature, which is a function of temperature and pressure. Although strictly these surfaces are not model levels, they are referred to as such in the remainder of this report for the sake of simplicity. The global model is run every 12 hr from initial conditions valid at 00:00 and 12:00.

3. The SAFHQ Aircraft Icing Algorithm

In order to aid the forecasting of aircraft icing over Sweden, SAFHQ has recently started applying a new icing algorithm to the output of a configuration of the European HIgh Resolution Limited Area Model (HIRLAM) [3]. The configuration of this NWP model used by SAFHQ is similar to the mesoscale model. Both are high resolution limited area models with rotated pole latitude-longitude grids and similar pressure based vertical coordinate

systems. SAFHQ's configuration of HIRLAM, having adjacent gridpoints approximately 12 nm apart and 31 levels in the vertical, also has a similar resolution to the mesoscale model. However, it should be noted that, differences between the HIRLAM and the mesoscale model mean that the conclusions of the assessment of the forecasts from the mesoscale model presented in this document do not necessarily also apply to the SAFHQ icing forecasts.

The output of the new SAFHQ aircraft icing algorithm is in terms of two types of icing intensity index. Forecast fields of each index are derived via the following procedures, using forecast fields of liquid water content (LWC in g/kg), temperature (T in °C) and vertical wind component (w in cm/s) outputted from an NWP model:

(i) Index *A*

1) Preliminary values of *A* are calculated as below for all gridpoints in the forecast fields.

$$A = 5 + \ln(\text{LWC})$$

2) The value of *A* is then set to 0 for all gridpoints except where...

stratiform cloud exists and $0 > T > -15$
convective cloud exists and $0 > T > -25$

3) Finally, the nonzero values of *A* are all rounded to the nearest integer. Since typical values of LWC do not exceed 1 g/kg and negligible values of LWC are not considered, ln(LWC) always lies between –5 and 0. Hence *A* ranges between 0 and 5, 0 indicating no risk of icing and 5 indicating a risk of severe icing.

(ii) Index *B*

1) The value of *B* is set to 0 for all gridpoints except where...

LWC > 0
$T < 0$

2) Where *B* has not already been set to 0 it is calculated from Table 1 below.

Table 1 - The derivation of *B* from *A* and *w*

	Descent	**Ascent**			
w	< 0	0 –10	10 – 20	20 – 30	> 30
B	*A* –1	*A* + 1	*A* + 2	*A* + 3	*A* + 4

3) Finally, negative values of *B* are all set to 0. Hence *B* ranges between 0 and 9, 0 indicating no risk of icing and 9 indicating a risk of severe icing.

4. Basic Assessment Method

The assessment of the aircraft icing forecasts involved a database of observational data supplied by Qinetiq. The database consists of pilot reports of encounters of RAF aircraft with icing that occurred during 2000 and 2001. All of these encounters occurred at altitudes below 25,000 feet (ft) and most occurred during flights to or from the UK, often over the UK itself. The altitudes of the encounters were generally reported to the nearest 500 ft but the precision of the locations of the encounters in the horizontal varied from an exact bearing and distance from a named navigation beacon to a description of a geographical area a few hundred nautical miles in length. The days, departure points and destinations of flights on which no icing was encountered were also included in the database but more precise information on flight paths and times was not given for these entries.

28 of the entries in the Qinetiq database corresponded to icing encounters that took place over the UK at well defined times and locations. The times and locations of these 28 icing encounters were compared with corresponding forecast fields of SLWC and the SAFHQ algorithm's icing intensity indices *A* and *B* covering the UK and environs from the mesoscale model. The SLWC fields were derived by setting SLWC equal to LWC where $T < 0$ and SLWC equal to zero where $T > 0$. The *A* and *B* fields were calculated from the SLWC fields and the corresponding *w* fields from the model. The method used was equivalent to that described in Section 3 with the exception that, for simplicity, index *A* was calculated as $A = 5 + \ln(\text{SLWC})$ and the lower bounds on temperature for *A* to be nonzero were not applied. Since the mesoscale model is run every 6 hr, two sets of forecast fields for each of the 28 icing encounters were available - one within the +6 to +12 hr range and one within the +12 to +18 hr range. Hence, a total of 56 forecasts from the mesoscale model were investigated.

Forecast fields of SLWC, *A* and *B* covering the UK and environs from the global model were derived as described in the previous paragraph for the mesoscale model. However, of the 28 icing encounters considered in the assessment of forecasts from the mesoscale model, only 21 were considered in the assessment of forecasts from the global model. This was because time constraints made it impractical to complete enough forecast runs of version 5.2 of the global model for all 28 encounters to be considered. Time constraints were less of an issue for the assessment of the mesoscale forecasts as all of the necessary UM version 4.5 output existed prior to the study as part of the output of the Met Office's operational suite of NWP models. Since the global model is run every 12 hr, only one set of forecast fields within the +6 to +18 hr range was available for each of the 21 icing encounters considered. Hence, a total of 21 forecasts from the global model were investigated.

5. General Observations

Figure 1 overleaf shows maps of a four model level thick slice of the forecast fields for one of the mesoscale model forecasts. Figure 2 overleaf shows maps

of the equivalent global model forecast fields. Figures illustrating these particular forecasts have been included as they illustrate typical features of the other forecasts. The comments that follow in this section therefore apply not only to the example forecasts shown in the figures but also to all of the other forecasts considered in this study as well. The fields illustrated in Figures 1 and 2 were calculated using +11 hr data from the 12:00 run of the models on 2nd February 2001. The validity time for the fields is therefore 23:00 on the same day. The Qinetiq database includes an icing encounter that took place at 23:05. The location of the icing encounter in the horizontal is marked on the maps in both figures with an orange cross. The location of the encounter in the vertical is marked by a horizontal dashed line running across the centre of each figure. SLWC, *A* and *B* fields for the model levels immediately above and below the altitude of the encounter are shown, as well as the same data for the model levels immediately above and below them respectively. Each column of maps corresponds to one of the types of fields while each row of maps corresponds to a model level. The approximate altitude of each model level in the vicinity of the icing encounter can be found on the extreme left of each figure. Neither of the models have model levels following surfaces of constant altitude so the altitudes given are not necessarily applicable to areas of the fields far from the icing encounter. Keys to the colours used in each column of maps can be found at the bottom of each figure. Note that the black areas of the maps in the left-hand columns correspond to gridpoints having any value of SLWC above 0.5. Similarly, the black areas in the right–hand columns correspond to gridpoints having values of index *B* of 6, 7, 8 or 9. There are no black areas in the middle columns of maps because index *A* cannot have a value greater than 5.

The coarser horizontal resolution of the global model becomes evident when the maps in Figures 1 and 2 are compared in terms of the areas where aircraft icing might occur, identified by gridpoints with nonzero SLWC, *A* and *B*. Such areas are evident on all the model levels shown in both figures. There is a degree of correspondence between the areas of potential icing forecast by the two models but there are also clear differences in their location, extent and shape. Such differences are evident in the case of the area of potential icing located in the southwest corner of the maps for the model levels immediately above and below the icing encounter.

Comparing the maps within each row of each figure, it can be seen that areas of potential icing identified by nonzero SLWC gridpoints are generally larger than the areas identified by nonzero *B* gridpoints, which are larger than those identified by nonzero *A* gridpoints.

Figure 1 – Mesoscale model +11 hr forecast fields of SLWC, *A* and *B* corresponding to an icing encounter occurring at 23:05 on 2nd February 2001

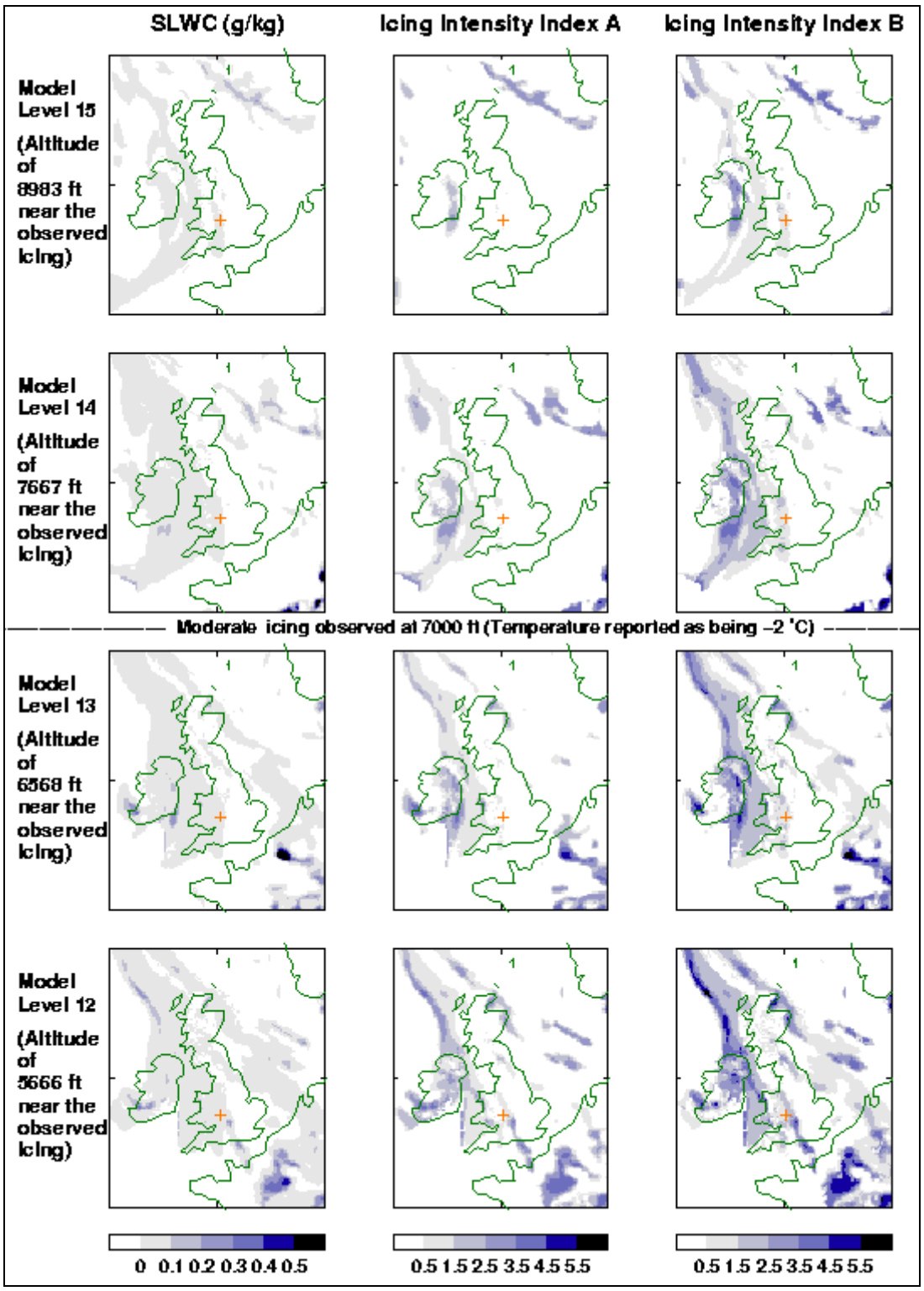
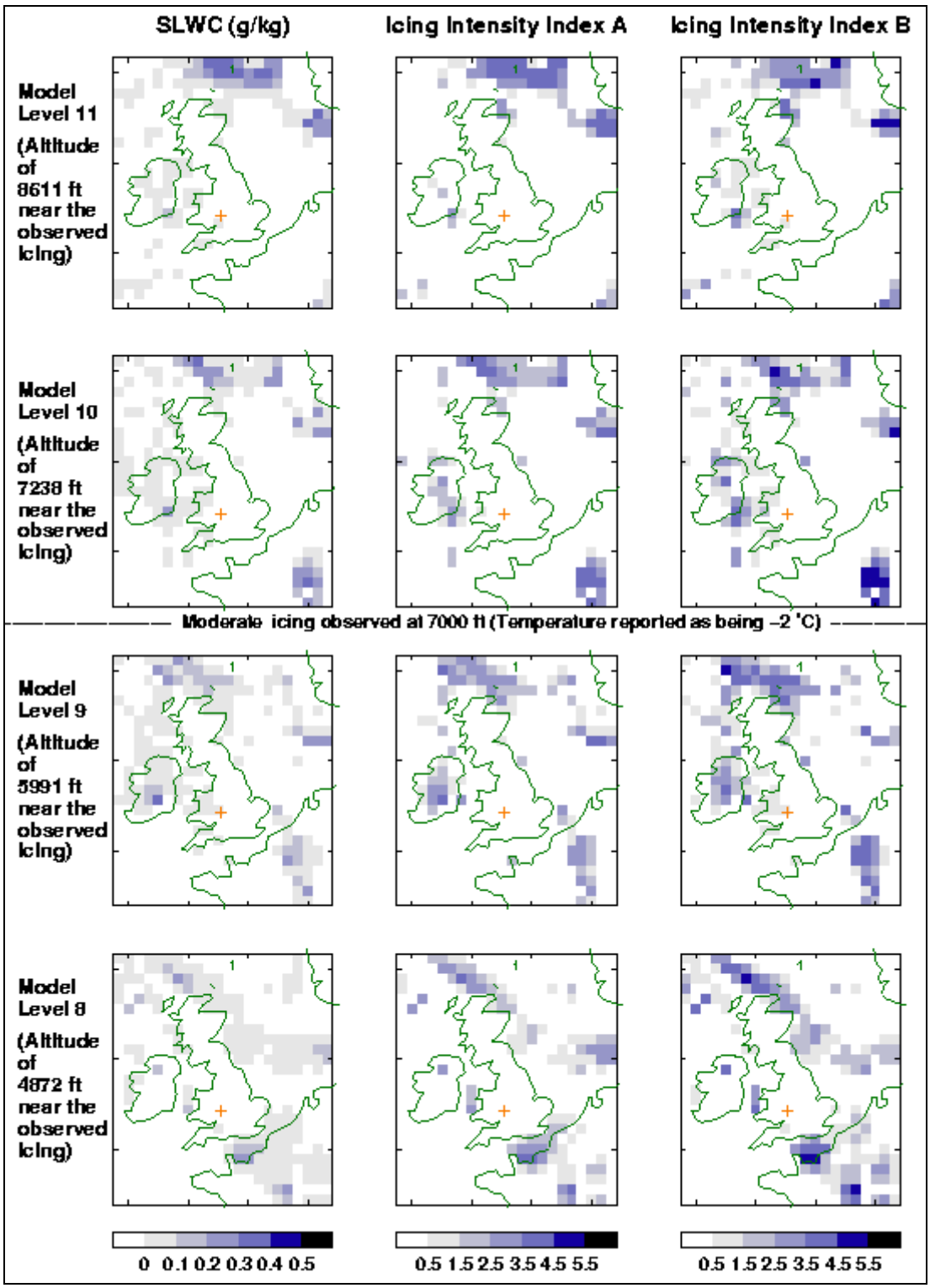


Figure 2 – Global model +11 hr forecast fields of SLWC, *A* and *B* corresponding to an icing encounter occurring at 23:05 on 2nd February 2001



Comparing the middle and right-hand columns of each figure, it can also be seen that a typical difference between the *A* and *B* fields is that the *B* fields have more structure to them than the *A* fields. Index *B* appears to be more capable of identifying small areas of moderate icing intensity within much larger areas of less significant icing intensity than index *A*. This is much more

noticeable in the forecast fields from the mesoscale model than in those from the global model. As an initial assessment of icing forecasts from the Met Office global and UK mesoscale NWP model the remainder of this study ignores the structure in the forecast fields of SLWC, *A* and *B*. Although SLWC, *A* and *B* all took numerous different nonzero positive values at different times and locations, and the frequency distributions of these values are perhaps different for the global and mesoscale model forecasts, the following assessments of hit rate and false alarm rate made no distinction between these nonzero values and considered only the distinction between zero and nonzero values. Any significance in the variation between nonzero values of SLWC, *A* and *B* or the differences in their frequency distributions was not considered and all forecast nonzero values of SLWC, *A* and *B* were considered to be predictors of a non-negligible risk of icing.

6. Hit Rate Assessments

It was noted that, because of the finite resolutions of the models and the relatively small number of icing encounters considered, in no forecast did a gridpoint with nonzero SLWC actually coincide exactly with the location of an icing encounter. The left-hand column of maps in Figure 1 might appear to suggest otherwise but it should be noted that each gridpoint with nonzero SLWC is represented in the maps by a coloured block centred on the location of the gridpoint rather then by a more representative, but less visible, coloured dot at the location of the gridpoint. A hit rate for forecasts could have been derived by defining a "hit" to have occurred only when potential icing conditions were forecast to occur at the nearest gridpoint to an icing encounter. However, such an analysis would have yielded no information about forecasts which forecast icing near, but not at the nearest gridpoint to, an icing encounter. In addition, due to the different resolutions of the models, "hits" from the 56 mesoscale model forecasts and "hits" from the 21 global model forecasts would not be equivalent in terms of the actual proximity of icing encounters to forecast areas of potential icing. This would render a comparison of the mesoscale model hit rate and the global model hit rate meaningless. In view of all this, hit rate was not assessed in terms of a literal hit rate but instead assessed in terms of the horizontal separations between the icing encounters and the locations of the nearest gridpoints with nonzero values of SLWC, *A* and *B* and the altitudes of these gridpoints relative to their corresponding icing encounters. A similar approach was used by R.W. Lunnon in an assessment of LWC forecasts for North Sea helicopter operations [4].

The results of the hit rate assessments for the mesoscale model forecasts and the global model forecasts are illustrated separately in Figures 3 and 4 overleaf. The vertical dashed lines in Figures 3 and 4 divide the forecasts into different categories according to the horizontal separation between each icing encounter and the location of the nearest forecast icing. The numbers of forecasts in each category are given in Table 2 below, on which a discussion of

the horizontal separations between the icing encounters and the locations of their nearest icing predictors is based. The horizontal dashed lines in the figures show the values of the mean, the mean plus twice the standard deviation, and the mean minus twice the standard deviation of the frequency distributions of the altitudes of the nearest predictors relative to their corresponding icing encounters. These numbers are given in Table 3 overleaf, on which a discussion of the altitudes of the nearest predictors to their corresponding icing encounters is based. This discussion and that of the horizontal separations is summarised at the end of the section.

(i) Horizontal separations between encounters and their nearest predictors

Table 2 - The frequency distributions of the horizontal separations between icing encounters and their nearest predictors

Horizontal Separation (nm)	Mesoscale Model Forecasts (%)			Global Model Forecasts (%)		
	Nonzero SLWC	Nonzero A	Nonzero B	Nonzero SLWC	Nonzero A	Nonzero B
1 to 4	73	51	62	0	0	0
4 to 16	14	30	25	52	29	38
16 to 64	11	14	11	48	38	62
64 to 100	2	5	2	0	33	0
1 to 16	87	81	87	52	29	38
1 to 64	98	95	98	100	67	100

No icing was forecast within 1 nm of an icing encounter in any of the mesoscale model forecasts and none was forecast within 4 nm of an encounter in any of the global model forecasts. This is unsurprising given the different finite horizontal resolutions of the models. The finer horizontal resolution of the mesoscale model gave it a clear advantage over the global model in terms of minimising the horizontal separations between the icing encounters and their nearest predictors. At least half of the mesoscale model forecasts of any type forecast icing within 4 nm of an encounter. In addition, at least 80% of the mesoscale model forecasts of any type forecast icing within 16 nm of an encounter while the equivalent maximum percentage for the three types of global model forecasts was around 50%. This superiority of the mesoscale model was less in the case of more poorly forecast encounters where the horizontal separations between the icing encounters and their nearest predictors exceeded 64 nm, approximately twice the separation between adjacent global model gridpoints. Although only about two thirds of the nonzero *A* global model forecasts forecast icing within 64 nm of an icing encounter, almost all of each of the other five types of forecast forecast icing within 64 nm of an encounter. In terms of the horizontal separations between the icing encounters and their nearest predictors then, the worst mesoscale model forecasts were as poor as the worst nonzero SLWC and nonzero *B* global model forecasts but not the nonzero *A* global model forecasts, which were poorer.

For the mesoscale model forecasts, the three frequency distributions of horizontal separations between icing encounters and the nearest forecast icing were similar for separations greater than 16 nm , especially in the case of

nonzero SLWC and nonzero *B*. However, the distributions were very different for separations less than this. The ratio of separations less than 4 nm to those between 4 and 16 nm was greatest for nonzero SLWC and least for nonzero *A* for the mesoscale forecasts. This comparison suggests that, in terms of the distributions of horizontal separations, nonzero SLWC performed better than nonzero *B*, which performed better than *A*, as a predictor of aircraft icing. The actual distribution for a certain predictor for horizontal separations approximately equal to or less than 6 nm, the approximate horizontal resolution of the mesoscale model, was undoubtedly dominated by the distances between the encounters and their nearest gridpoints. However, since the set of these distances was the same for each distribution, this did not invalidate this comparison of the three distributions.

The same argument holds true for a comparison of the three distributions of horizontal separations for the global model forecasts, although in this case the approximate horizontal resolution of the model is 32 nm. The percentage of separations less than 16 nm was greatest for nonzero SLWC and least for nonzero *A*. The inferiority of nonzero *A* as an icing predictor is also apparent for the larger horizontal separations where the nonzero *A* distribution for the global model forecasts was the only one of the six distributions summarised in Table 2 to have more than 5% of the separations greater than 64 nm. As for the mesoscale model forecasts, nonzero SLWC performed better than nonzero *B*, which performed better than *A*, in terms of the distributions of horizontal separations.

(ii) Altitudes of nearest predictors relative to their corresponding encounters

Table 3 – Statistics on the frequency distributions of the altitudes of the nearest predictors relative to their corresponding icing encounters

Statistic (Units are 1000 ft)	Mesoscale Model Forecasts			Global Model Forecasts		
	Nonzero SLWC	Nonzero A	Nonzero B	Nonzero SLWC	Nonzero A	Nonzero B
Mean	-2.5	-3.0	-2.8	0.8	0.8	1.1
SD	2.6	3.0	3.0	6.1	6.2	5.7
Mean+2SD	2.7	3.1	3.1	13.0	13.3	12.4
Mean-2SD	-7.7	-9.1	-8.8	-11.3	-11.6	-10.3

Figure 3 - The locations of the nearest predictors relative to their corresponding icing encounters for the mesoscale model forecasts

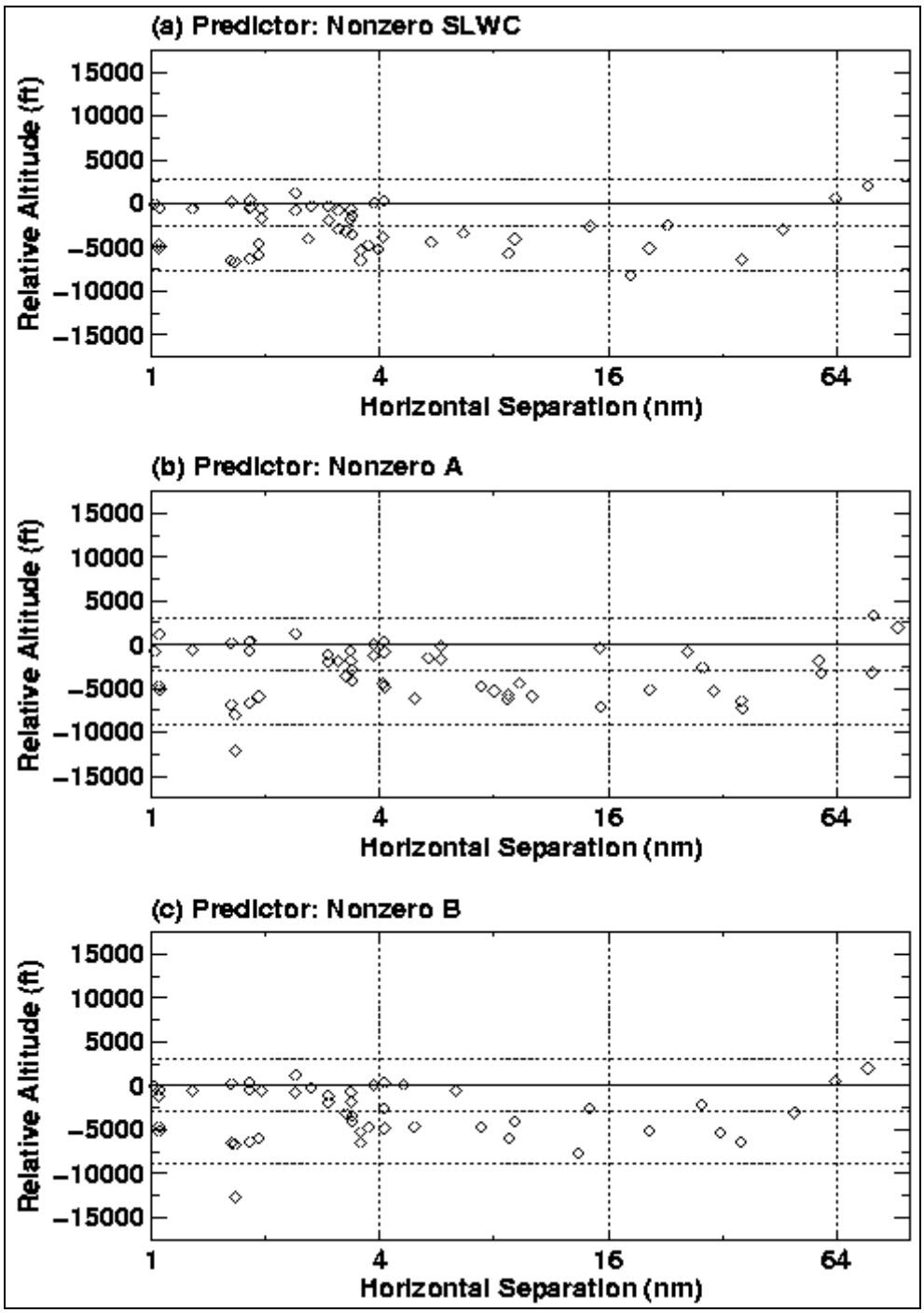
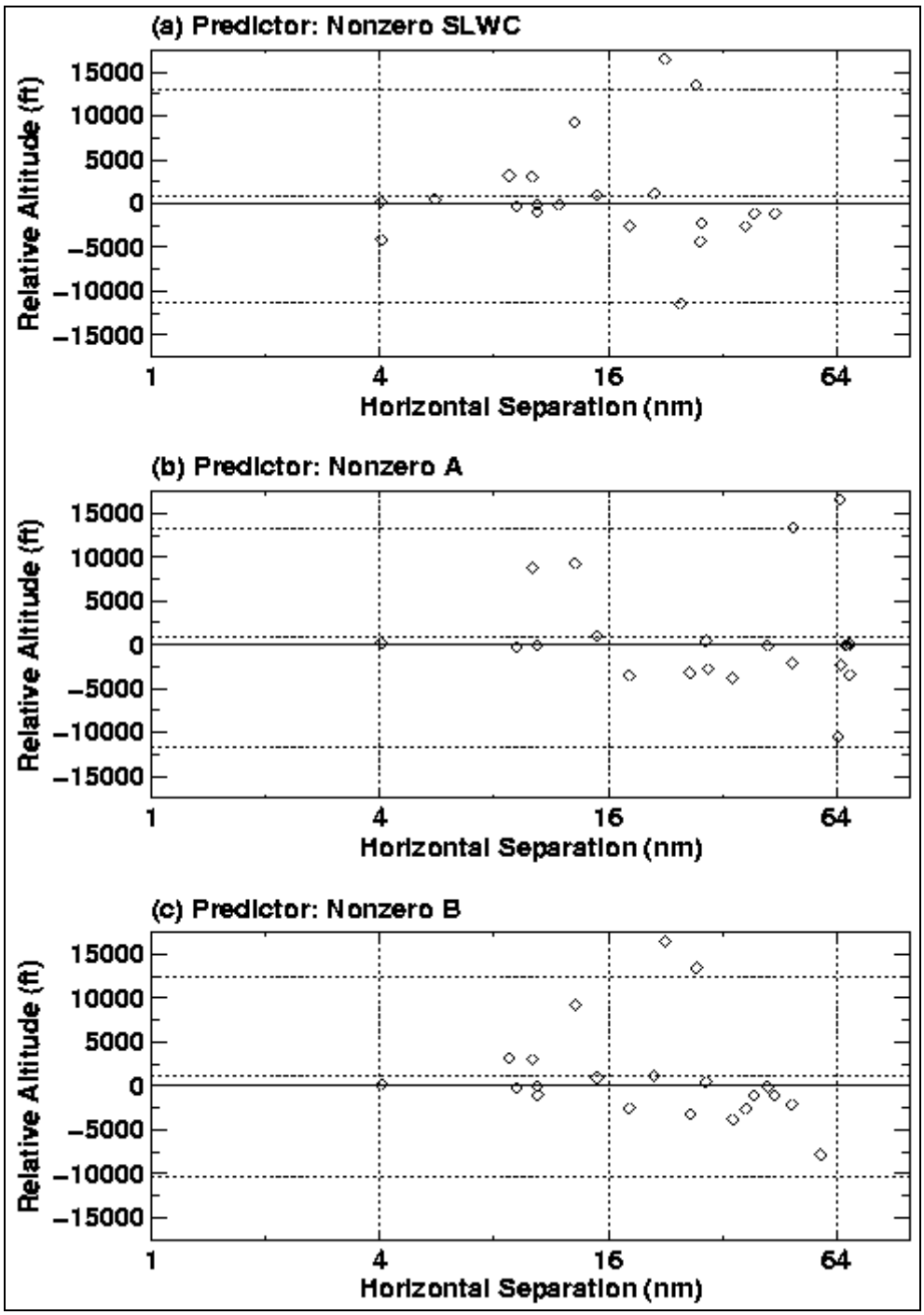


Figure 4 - The locations of the nearest predictors relative to their corresponding icing encounters for the global model forecasts



The frequency distributions of the altitudes of the nearest predictors relative to their corresponding icing encounters differed significantly between the mesoscale and global model forecasts. On average the mesoscale model forecast the nearest icing to an encounter to be about 2500 to 3000 ft below it, while the average relative altitude for the global model forecasts of a

particular type of predictor was either 800 or 1100 ft. Although the relative altitudes of the global model forecasts lacked the large negative bias of those of the mesoscale model forecasts, the standard deviations of the altitudes of the global model forecasts were about twice those of the mesoscale model forecasts. If each distribution of relative altitudes is assumed to be a sample of a normal distribution then it can be expected that 95% of global model forecasts will forecast the nearest icing to an encounter to be in a layer between about 11,000 ft below and 13,000 ft above the encounter. The equivalent layer for the mesoscale model forecasts is much shallower, between about 8000 ft below to about 3000 ft above the encounter. In terms of the altitudes of the nearest predictors relative to their corresponding icing encounters then, the mesoscale model forecasts appear to perform better than the global model forecasts.

Of the three types of mesoscale model forecast, the nonzero SLWC forecasts displayed the least bias towards underestimating the altitudes of the icing encounters. These forecasts also had the narrowest distribution of altitudes of the nearest predictors relative to their corresponding icing encounters. The distributions for the nonzero *A* and nonzero *B* forecasts were equally broad and both types of forecast displayed a similar degree of underestimation of the altitudes of the icing encounters. In terms of the altitudes of the nearest predictors relative to their corresponding icing encounters then, of the mesoscale model forecasts, the nonzero SLWC forecasts appear to perform better than nonzero *A* and nonzero *B* forecasts, which performed equally well.

If a normal distribution is assumed for the distribution of relative altitudes for the nonzero *B* global model forecasts then it can be expected that 95% of these forecasts will forecast the nearest icing to an encounter to be in a layer between about 10,000 ft below and 12,000 ft above the encounter. This layer falls entirely within the deeper and very similar equivalent layers for the nonzero SLWC and nonzero *A* forecasts, which are both between relative altitudes of about –11,000 and 13,000 ft. This suggests that, in terms of the altitudes of the nearest predictors relative to their corresponding icing encounters, of the global model forecasts, the nonzero *B* forecasts perform better than the nonzero SLWC and nonzero *A* forecasts, which performed equally well.

(iii) Summary of the results of the hit rate assessments

In general, the mesoscale model forecasts appeared to forecast potential icing conditions closer to the icing encounters, both in the horizontal and the vertical, than the global model forecasts. At least 80% of the mesoscale model forecasts of any type forecast icing within 16 nm of an icing encounter and 95% forecast icing in a layer between about 8000 ft below and about 3000 ft above an encounter. In contrast, a maximum of only about half of the global model forecasts of any type forecast icing within 16 nm of an icing encounter and 95% forecast icing in a layer between about 11,000 ft below and about 13,000 ft above an encounter. Therefore, it can be said that the mesoscale model forecasts have a higher hit rate than the global model forecasts.

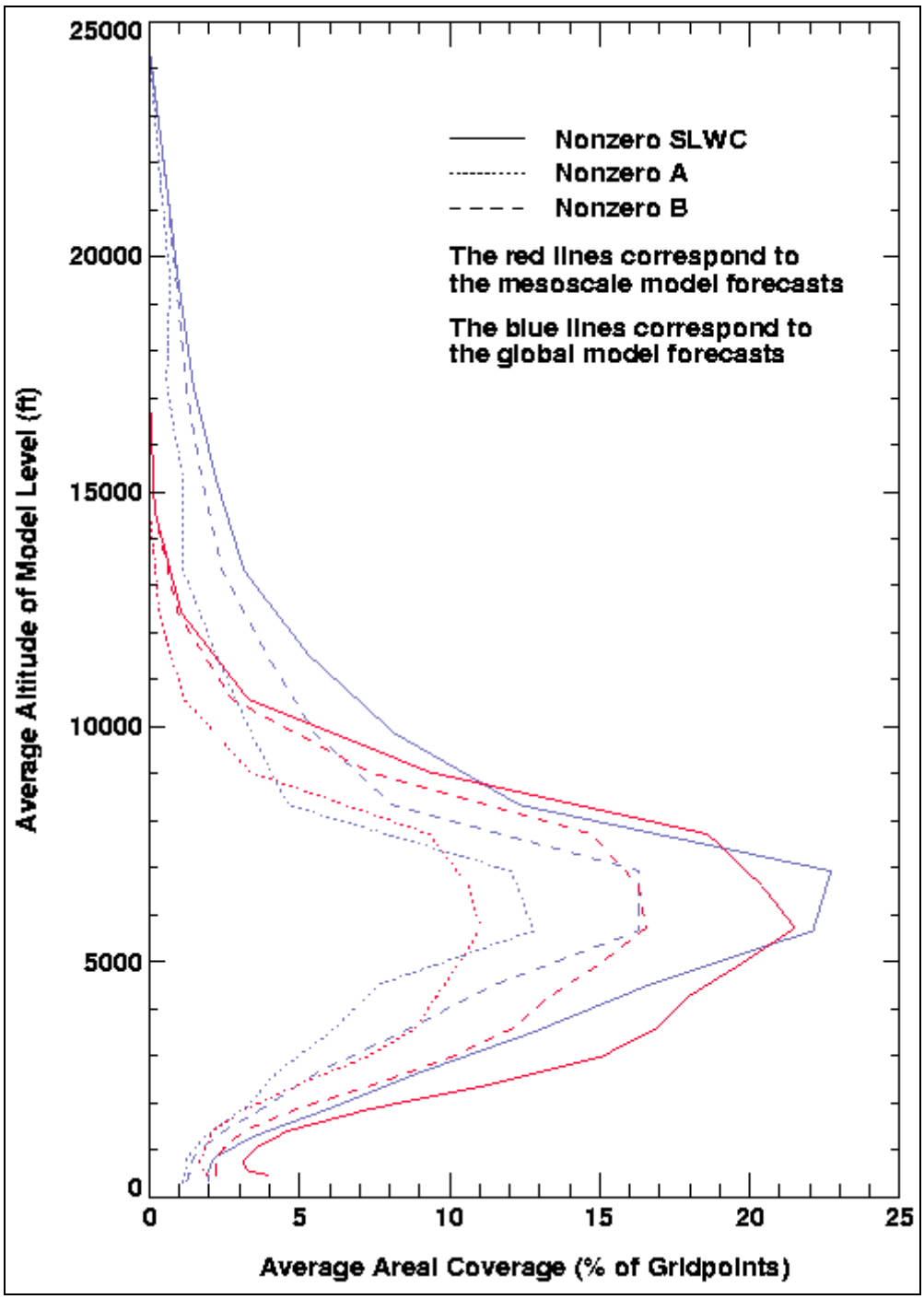
Of the three types of mesoscale model forecasts, the nonzero SLWC forecasts generally forecast potential icing conditions closest to the icing encounters, both in the horizontal and the vertical. Therefore, it can be said that the nonzero SLWC forecasts have a higher hit rate than the nonzero *A* and nonzero *B* forecasts. Both the nonzero *A* and nonzero *B* forecasts performed equally well in terms of forecasting the altitudes of icing encounters but the nonzero *B* forecasts forecast potential icing conditions closer to the icing encounters in the horizontal. Therefore, it can be said that the nonzero *B* forecasts have a higher hit rate than the nonzero *A* forecasts.

Of the three types of global model forecasts, the nonzero SLWC forecasts generally forecast potential icing conditions closest to the icing encounters in the horizontal but the nonzero *B* forecasts generally forecast icing conditions closest to the encounters in the vertical. If the accuracy of the forecasts in terms of the location of potential icing conditions in the horizontal is given priority over accuracy in terms of the location of icing conditions in the vertical then it can be said that the nonzero SLWC forecasts have a higher hit rate than the nonzero *B* forecasts. Both types of forecast have a higher hit rate than the nonzero *A* forecasts, which performed worst in terms of forecasting the locations of icing encounters in the horizontal.

7. False Alarm Rate Assessments

As mentioned in Section 5, a comparison of the maps within each row in Figures 1 and 2 reveals that the areas of potential icing identified by nonzero SLWC gridpoints are larger than the areas identified by nonzero *B* gridpoints, which are larger than those identified by nonzero *A* gridpoints, in both the mesoscale and global model forecasts. The fact that the extents of the nonzero SLWC areas are greater than the areas covered by the other two types of predictor is unsurprising. The definitions of *A* and *B* ensure that *A* and *B* can only be, but aren't constrained to be, nonzero at gridpoints where SLWC is nonzero. The relative areal coverages of the three icing predictors displayed in Figures 1 and 2 were found to be typical of all of the mesoscale model and global model forecast fields respectively, as suggested by Figure 5 overleaf.

Figure 5 - The areal coverage of icing predictors averaged over all forecast fields at each model level



Unfortunately, it is not known how much of the extent of the nonzero SLWC areas in excess of the nonzero *B* areas and nonzero *B* areas in excess of the nonzero *A* areas were erroneously indicating icing conditions in either the mesoscale or the global model forecasts. This is because, while the Qinetiq database contained sufficient information to define 28 point locations over the

UK where there was a non-negligible icing risk, it contained insufficient information on the absence of icing conditions to allow reliable estimates of the extents of areas of non-negligible icing risk to be made.

However, as nonzero *A* predictors occupy less gridpoints than each of the other two types of predictor in both types of forecasts, it would seem possible that the number of nonzero *A* gridpoints located in areas in which there was actually no risk of icing is smaller than for each of the other two types of predictor. Hence the false alarm rate for nonzero *A* is possibly less than for the other two types of predictor in the cases of both the mesoscale and global model forecasts. Similarly, as nonzero *B* predictors occupy less gridpoints than nonzero SLWC predictors, the false alarm rate for nonzero *B* is possibly less than that for nonzero SLWC. However, Figure 5 suggests that the false alarm rates for nonzero *B* and nonzero SLWC may not be greatly different at altitudes above about 12,000 ft and 17,000 ft for the mesoscale and global model forecasts respectively. It also suggests that, in forecasts from both models, the false alarm rates for nonzero *B* and nonzero *A* may be least different near the ground.

Comparing the red and blue lines in Figure 5, it can be seen that, for a given predictor, the percentage of gridpoints for which potential icing conditions were forecast was similar for the mesoscale model forecasts and the global model forecasts between about 5000 and 10,000 ft. It is therefore possible that, for a given predictor, the forecasts from both models had similar false alarm rates in this layer. Above this layer each predictor in the global model forecasts predicted potential icing conditions over a larger area than the corresponding mesoscale model forecasts on average. This suggests a greater false alarm rate for the global model forecasts than for the mesoscale model forecasts above 10,000 ft. Below the layer each predictor in the mesoscale model forecasts predicted potential icing conditions over a larger area than the corresponding global model forecasts on average. This suggests a greater false alarm rate for the mesoscale model forecasts than for the global model forecasts below 5000 ft.

8. Conclusions

There is a degree of correspondence between the areas of potential icing forecast by the mesoscale and global models but there are also clear differences in their location, extent and shape. These differences appear to ensure that forecasts from the two models perform differently in terms of both hit rate and false alarm rate. At least 80% of the mesoscale model forecasts of any type forecast icing within 16 nm of an icing encounter and 95% forecast icing in a layer between about 8000 ft below and about 3000 ft above an encounter. In contrast, a maximum of only about half of the global model forecasts of any type forecast icing within 16 nm of an icing encounter and 95% forecast icing in a layer between about 11,000 ft below and about 13,000 ft above an encounter. The mesoscale model forecasts therefore have a higher hit rate than the global model forecasts. Above 5000 ft they also have a possible false alarm rate approximately equal to or less then that of the

global model forecasts. The results of this study would therefore seem to suggest that the mesoscale model should be used in preference to the global model to forecast aircraft icing over the UK in the portion of the atmosphere above 5000 ft. The situation is less clear for the portion of the atmosphere below 5000 ft, where it is possible that the mesoscale model forecasts have a greater false alarm rate than the global model forecasts.

Of the three different types of icing predictor, nonzero SLWC appeared to perform best and nonzero *A* appeared to perform worst in terms of hit rate in both the mesoscale and global model forecasts. However, it is possible that in terms of false alarm rate, nonzero SLWC performed worst and nonzero *A* performed best. Much more about the relative performance of the predictors cannot be concluded without a more thorough study.

9. Recommendations for Further Work

It is the opinion of the author that if a further assessment of +6 to +18 hr aircraft icing forecasts from the mesoscale and global models is undertaken then the following should be included if possible.

(i) The use of observations taken after summer 2002

As mentioned in Section 2, at the time of writing, May 2002, version 5.2 of the UM is about to replace version 4.5 as the source of all of the Met Office's operational mesoscale and global NWP forecasts. The assessment of the mesoscale model forecasts presented here uses output from forecast runs of the mesoscale NWP configuration of version 4.5 forced with boundary conditions from the global NWP configuration of version 4.5. The assessment of the global model uses output from forecast runs of the global NWP configuration of version 5.2 initialised with initial conditions from version 4.5 runs. The results of neither of these assessments quite reflects the performance of the Met Office operational NWP models to be expected after version 5.2 of the UM becomes the standard operational version in summer 2002. If pilot reports of icing encounters occurring after summer 2002 became available a further study could use them to give a better assessment of the potential capability of the Met Office operational NWP models to forecast aircraft icing.

(ii) A rigorous false alarm rate assessment

The false alarm rate assessment described in this report is not rigorous and the conclusions in Section 8 should be considered with this in mind. A thorough assessment of the forecasts should include a more rigorous false alarm rate assessment, ideally based on pilot reports of well defined times and locations when and where icing was not encountered. Such observations were not available for this study. A rigorous false alarm rate assessment would perhaps allow more quantitative conclusions to be drawn about the quality of the different types of aircraft icing forecast.

(iii) An analysis considering the entire distribution of values of SLWC, *A* and *B*

As mentioned in Section 5, any significance in the variation between nonzero values of SLWC, *A* and *B* was not considered in this assessment. A further study should examine how hit rate and false alarm rate vary with threshold values of SLWC, *A* and *B* chosen to indicate the presence of icing conditions. This assessment only considers one such threshold - zero. In addition, any differences between the frequency distributions of the values of SLWC, *A* and *B* for the global and mesoscale model forecasts should be investigated. If more pilot reports were available of a range of severities of icing encounter a further study could also investigate the relationship between the severity of an observed icing encounter and the magnitudes of the nearest forecast nonzero values of SLWC, *A* and *B*.

The author also recommends that the following less important items be considered for inclusion in a further study.

(iv) An investigation into the variation of hit rates with altitude

In Section 7, this report briefly comments on the variation of possible false alarm rates with altitude. A further study could also investigate the variation of hit rates with altitude. Such an investigation would benefit from more pilot reports of icing encounters occurring at a range of altitudes being available.

(v) The use of observations taken outside the UK

This study has only made use of pilot reports of icing encounters that have taken place over the UK. It is not known if the conclusions in Section 8 apply to icing forecasts made for regions outside the UK. A further study could make use of observations taken outside the UK to establish this. A thorough assessment of the worldwide performance of icing forecasts from the mesoscale model would also necessitate the use of mesoscale domains centred on regions other than the British Isles.

(vi) An investigation into errors in LWC and *T* forecast fields

For particularly poor forecasts the relevant LWC and *T* fields could be compared with observations in order to determine the source of the error in the forecast SLWC fields. The Qinetiq database contains temperature observations coincident with the icing encounters so no more observations would be needed to investigate the accuracy of *T* forecast fields. However, a whole new set of observations would be needed to investigate the accuracy of LWC forecast fields as the Qinetiq database contains no LWC observations. Furthermore, a database that included LWC observations would probably need to be sourced from a research aircraft as it is generally only research aircraft that take LWC observations. A study including LWC observations from a research aircraft might be similar to R. Brown's March 2000 evaluation of the applicability of SLWC forecasts to the problem of forecasting aircraft icing [5].

(vii) Consideration of the time differences between forecasts and observations

This study took no account of the fact that, for any icing encounter not occurring "on the hour", there is a difference between the time that the icing was observed and the validity times of the forecasts that the encounter was compared with. A further study could attempt to take account of these time differences.

It should be emphasised that the implementation of many of these seven recommendations would benefit from more pilot reports of well defined times and locations when and where icing was and was not encountered. The availability of such observations and the time available to carry out this study have been the main factors that have limited the conclusions of this report. If time is made available for a further study, the most will be gained from that time if more relevant observational data is also available. At the time of writing the most promising source of such data appears to be British Airways as it is possible that data from icing sensors on the airline's Boeing 747 aircraft may become available to the Met Office in the future.

10. Acknowledgements

R.W. Lunnon, P. Agnew, S. Heyman, G. Greed, S. Milton and J. Swift of the Met Office all provided guidance to the author at various stages of the project. Thanks must also go to R.W. Gent and G.P. Burkitt of Qinetiq for supplying the database of pilot reports.

11. References

[1] - Met Office Operational Numerical Modelling Webpage (http://www.metoffice.com/research/nwp/numerical/operational/index.html)

[2] - Olofsson, B., Olsson, E., and Martensson, T., "A new algorithm to estimate aircraft icing in the HIRLAM model", American Meteorological Society Tenth Conference on Aviation, Range and Aerospace Meteorology of May 2002.

[3] – The International HIRLAM Project Webpage (http://www.knmi.nl/hirlam/)

[4] - Lunnon, R.W., "Numerical forecasting of liquid water content to assess airframe icing risk", American Meteorological Society Fourth International Conference on Aviation Weather Systems of June 1991.

[5] – Brown, R., "An evaluation of the applicability of supercooled liquid water content forecasts to the forecasting of aircraft icing", Met Office Forecasting Research Technical Report Number 305 of March 2000.